

# DNA ベースの自己組織化ネットワーク探索ツールへの oxDNA の導入

竹口 葵衣 (指導教員：オベル加藤ナタナエル)

## 1 Introduction

DNA は一定条件下の混合により特定の形状へと組み上がる「自己組織化」という性質をもつが、これは近年「DNA オリガミ」と呼ばれる DNA を利用して様々なナノ構造を生成する技術に利用され、それらは主に医療分野において薬物分子の送達や体内の探索への応用が期待されている [1]。このような DNA 構造の設計は主にコンピュータソフトウェアによるが、DNA 構造の生成は非常に多くの条件に左右されるため、それらを直接設計することは困難である。Cazenille らにより、DNA の自己組織化における化学反応ネットワークを自動的に探索するツールが研究されている [2] もの、その出力結果は NUPACK による構造の予測にとどまり、実世界における生成可能性の検証が可能とはいえない。本研究では、このツールの出力にシミュレーションツールである oxDNA を導入し、生成可能な DNA 構造を得る方法を提案する。

## 2 データ構造と使用ツール

### 2.1 DNA 構造の化学反応ネットワーク

CRN (Chemical Reaction Network、化学反応ネットワーク) とは、物質をノード、物質間の化学反応の流れをエッジとするネットワークである。今回は各 DNA 構造とそれら同士が反応する過程をネットワークのグラフとして表現している。

### 2.2 列挙ツール “Peppercorn”

Peppercorn は、Badelt らによって開発された DNA 鎖置換反応の列挙ツールで、有限の初期複合体 (DNA 構造) のセットとそれらの相互作用のルールのセットが与えられた場合に、CRN を生成する「列挙」を行う [3]。

### 2.3 QD アルゴリズムと MAP-Elites

QD (Quality-Diversity) アルゴリズムは、解の集合から一定確率で選択した値を交叉・変異させ、それを評価した上で解集合に配置することを繰り返し、多様な最良解の集合を得るアルゴリズムである。[4] その一種である MAP-Elites は、ユーザが選んだ  $N$  個の変動次元 (対象の大きさやエネルギー等のパラメータ) が与えられたときに、それらの次元で定義される空間内の各点で最も最良な解 (エリート) を探索し、それらを座標平面 (グリッド) に配置するアルゴリズムである [2]。「子孫の性能・特徴を決定し、座標に配置する (座標にすでに親ゲノムがあれば、より解に適する方を配置する)」という特徴がある。

### 2.4 NUPACK と oxDNA

NUPACK は核酸の溶液中の熱力学的解析、核酸鎖の複合体の配列設計、核酸鎖の複合体の平衡特性の評価・表示・注釈などを提供するツールである [6]。一方、oxDNA は与えられた塩基配列について、付属の python コードを用いて初期構成と外力を設定して実行すると、実世界における反応過程および挙動をシミュレートし、3D イメージに変換可能な数値データ (反応時の分子の軌道、エネルギー、最終構成など) を提供するオープンソースソフトウェアである。[7]

### 2.5 DNA 自己組織化の自動探索ツール

Cazenille らは、Peppercorn と MAP-Elites を利用して、特定の DNA 鎖から CRN を検索し、DNA が様々な構造を形成する条件を特定する自動探索ツールを開発した。DNA 鎖のライブラリの内容や規模などを設定しプログラムを起動すると、CRN の各ノードにあたる構造と各エッジにあたる化学反応の流れがドメイン (一定の長さの DNA 鎖) 単位で出力され、それらを NUPACK で解析することで 3D イメージとして構造を確認することができる。その研究では、複数の規模が異なるライブラリを解析し、実在しうる構造のネットワークを探索して、それぞれを 1 つの点として平均構造サイズと反応タイプのエントロピーに基づいて座標 ( $x$ =反応数、 $y$ =構造数) に出力した (図 1)。

## 3 実験

本研究では、上記の探索ツールによって得られる L1 の一つのネットワークについて最適化を行い、得られたドメイン単位の各構造をランダムな塩基配列に翻訳した。使用 DNA 数は 30 鎖まで、生成可能な構造は 400 種類、反応の種類は 2000 種まで、シミュレーション時間  $\text{step} = 1,000,000 \mu\text{s}$  とし、各塩基配列を記述したファイルを入力して oxDNA を実行して 3D イメージおよびその構造の各ヌクレオチドの座標を取得した。また、各構造について、その座標による凸包 (各点を頂点とする多角形) と、各点を内包する直方体の体積を算出し、ヒストグラムを作成した。oxDNA の仕様では長さの単位が  $0.8518\text{nm}$  と定義されることに基づき、体積の単位はその 3 乗の  $0.61803476783\text{nm}^3$  とする。

## 4 結果

各構造から生成した塩基配列のファイルから、oxDNA によりシミュレーション結果として軌道・エネルギー・最終構成ファイルを取得でき、oxDNA に付属する python コードを用いて 3D イメージ化 (pdb 形式) することが

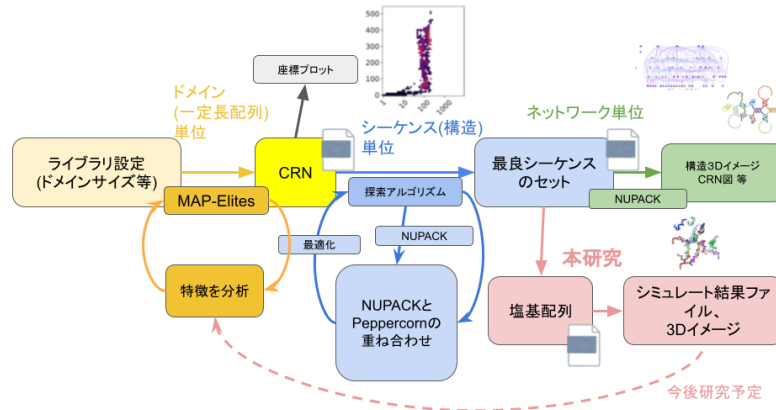


Figure 1: 探索ツールの基本的な動作。Cazenille らにより開発。[2] より引用。

できた (図 2)。また、各構造について凸包と直方体の体

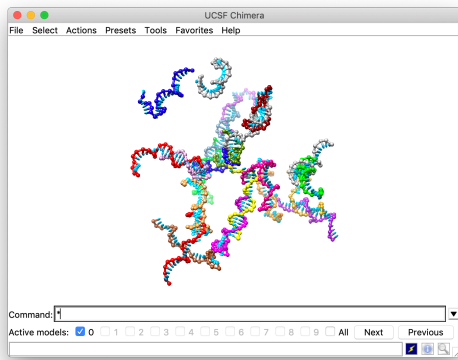


Figure 2: steps=1,000,000、ID が 2025 である構造の 3D イメージ。UCSF Chimera により表示。

積をヒストグラム化すると、どちらも偏りはあるが、その大きさは様々な値を示した。これにより、実験の結果様々な種類の構造が得られたことがわかる。一方、各構造に対してのシミュレーションにおける「DNA 鎖が結合した最大の数」は、予想された数 (=シミュレーションに使用された DNA 鎖の数) よりも少なかった。これは、すべての DNA 鎖が結合するにはシミュレーション時間は step = 1,000,000 では不足していることを表す。(図 3)

## 5 結論と議論

今回の研究において、Peppercorn および MAP-Elites アルゴリズムを用いた DNA 構造自動探索ツールにより発見された構造が oxDNA シミュレーションが可能であることを示した。これにより、NUPACK で予想された構造が作成可能かどうかをコンピュータ上で検証することが可能となった。今回の実験は生成に成功した一つのネットワークに対してのみ行われたが、先行研究では複数のネットワークにおける分析が行われた。今後の課題としては、生成に成功したすべてのネットワークにおいてシミュレーションを適用し、各結果に基づくライブラリの考察が必要であると考えられる。

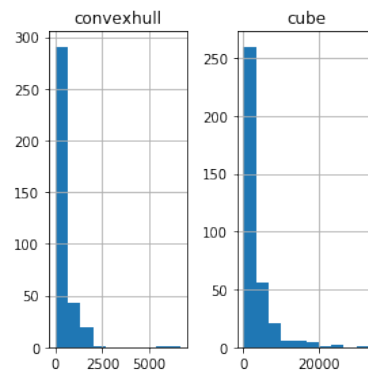


Figure 3: 凸包 (convexhull) と直方体 (cube) の体積のヒストグラム。横軸は体積 (単位:0.61803476783nm<sup>3</sup>)。凸包では約 2500 まで、直方体では約 10000 までの体積を持つ構造が多くを占めるが、範囲としてはそれぞれおよそ 6676、33427 となっており、規模の異なる多くの種類の構造が得られたことがわかる。

## 参考文献

- [1] Chou, L. Y. (2021). Design Verification as Foundation for Advancing DNA Nanotechnology Applications. *ACS Nano*, 15(6), 9222-9228
- [2] Cazenille, L., Baccouche, A., & Aubert-Kato, N. (2021). Automated exploration of DNA-based structure self-assembly networks. *J. R. Soc. Open Science*, 8(10), 210848.
- [3] S Badelt *et al.* A domain-level DNA strand displacement reaction enumerator allowing arbitrary non-pseudoknotted secondary structures. *J. R. Soc. Interface*, 17(167):20190866, 2020.
- [4] Mouret, J. B., & Clune, J. (2015). Illuminating search spaces by mapping elites. *arXiv preprint arXiv:1504.04909*.
- [5] Cully, A., Mouret, J. B., & Doncieux, S. (2021, July). Quality-diversity optimisation. *GECCO 2021*, 715-739.
- [6] Zadeh, *et al.* (2011). NUPACK: analysis and design of nucleic acid systems. *J. Comput. Chem.*, 32(1), 170-173.
- [7] P. Šulc, F. Romano, T. E. Ouldrige, L. Rovigatti, J. P. K. Doye, A. A. Louis (2012). Sequence-dependent thermodynamics of a coarse-grained DNA model. *J. Chem. Phys.* 137, 135101